Convex formulation and global optimization for multimodal active contour segmentation

Jonas De Vylder, Filip Rooms, Wilfried Philips
Department of Telecommunications and Information Processing,
IBBT - Image Processing and Interpretation
Ghent University
Email: Jonas.deVylder@telin.ugent.be

***Abstract*—Region based active contours have been proven useful in a wide range of applications. This method however hampers from the drawback that in only allows bimodal segmentation, i.e. foreground and background are two regions with approximately uniform intensity. In this paper we propose a new active contour with convex energy which allows multimodal foreground and background.**

Since Kass et al. [1] introduced the active contours, the framework has become a constant recurring topic in segmentation literature [2], [3], [4], [5], [6]. In the active contour framework, an initial contour is moved and deformed in order to minimize a specific energy function. This energy function should be minimal when the contour is delineating the object of interest, e.g. a leaf. Two main groups can be distinguished in the active contour framework: one group representing the active contour explicitly as a parameterized curve and a second group which represents the contour implicitly, e.g. using level-sets. In the first group, also called snakes, the contour generally converges towards edges in the image [1], [4]. The second group generally has an energy function based on region properties, such as variance of intensity of the enclosed segment [3], [7]. These level-set approaches have gained a lot of interest since they have some benefits over snakes. For example, they can easily change their topology, e.g. splitting a segment into multiple unconnected segments. Recently an active contour model has been proposed with a convex energy function, making it possible to define fast global optimizers [5], [6]. These global active contours have the benefit that their result no longer depends on the initialization.

In [8], Bresson et al. proposed an active contour with a convex energy function which combines edge information and region information. This method combines the original snake model [1] with the active contour model without edges [3]. This method however still hampers from the same constraint that the original Chan-Vese active contours has: it only allows bimodal segmentation, i.e. it only distinguishes between a dark background and a bright foreground, or vice versa. This is a reasonable assumption if foreground and background are only considered in a local area around the contour, such as sometimes happens with level-sets. But for a global optimizer where foreground and background are considered for the complete image, this becomes a drawback for a wide range of applications. In this paper we propose a new convex formulation for active contours which does allow foreground and background with multimodal intensities.

Jonas De Vylder is funded by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT).

This paper is arranged as follows. The next section provides a brief description of convex energy active contours. In section III our proposed algorithm is presented. Section IV elaborates on the optimization of the proposed convex energy function. Next, section V shows the results of our technique and is compared to the results from the bimodal convex active contour formulation. Section VI recapitulates and concludes.

I. Notations and Definitions

In the remaining of this paper we will use specific notations. To make sure all notations are clear, we briefly summarize the notations and symbols used in this work.

We will refer to an image, F in its vector notation, i.e. $\mathbf{f}(i*m+j) = F(i,j)$, where $m \times n$ is the dimension of the image. In a similar way we will represent the active contour in vector format, $\mathbf{u}$. If a pixel $U(i,j)$ is part of the segment, it will have a value above a certain threshold, all background pixels will have a value lower than the given threshold. Note that this is similar to level-sets. The way these contours are optimized however is different than with classical level-set active contours, as is explained in the next section. We will use image operators, i.e. gradient, divergence and Laplacian in combination with this vector notation, however the semantics of the image operators remains the same as if it was used on a matrix:

$$\begin{aligned} \nabla\mathbf{f}(t) &= \big(\nabla_x\mathbf{f}(t), \nabla_y\mathbf{f}(t)\big) \\ \nabla\cdot\mathbf{f}(t) &= \nabla_x\mathbf{f}(t) + \nabla_y\mathbf{f}(t) \\ \nabla^2\mathbf{f}(t) &= \nabla\cdot\nabla\mathbf{f}(t)) \end{aligned}$$

where

$$\begin{aligned} \nabla_x\mathbf{f}(i*m+j) &= F(i+1,j) - F(i,j) \\ \nabla_y\mathbf{f}(i*m+j) &= F(i,j+1) - F(i,j) \end{aligned}$$

Further we will use the following inner product and norm

notations:

$$\begin{aligned}\langle \mathbf{f}, \mathbf{g} \rangle &= \sum_{i=1}^{mn} f(i), g(i) \\ |\mathbf{f}|_{\mathbf{g}} &= \sum_{i=1}^{mn} g(i) \mid f(i) \mid\end{aligned}$$

If the weights $g(i) = 1$ for all i, then we will omit $\mathbf{g}$, since we assume this will not cause confusion, but will increase readability.

II. Convex Energy Active Contours

In [5] an active contour model was proposed which has global minimisers. This active contour is calculated by minimizing the following convex energy:

$$E[\mathbf{u}] = |\nabla \mathbf{u}| + \gamma \langle \mathbf{u}, \mathbf{r} \rangle \tag{1}$$

with

$$r(x) = (\mu_f - f(x))^2 - (\mu_b - f(x))^2 \tag{2}$$

Here $\mathbf{f}$ represents the intensity values in the image, μ_f and μ_b are respectively the mean intensity of the segment and the mean intensity of the background, i.e. every pixel not belonging to the segment. Note that this energy is convex, only if μ_f and μ_b are constant. If these values are not known in advance, they can be approximated by alternating between the following two steps: first fix μ_f and μ_b and minimize eq. (1), secondly update μ_f and μ_b. Chan et al. found that the steady state of the gradient flow corresponding to this energy, i.e.

$$\frac{d\mathbf{u}}{dt} = \nabla \cdot \frac{\nabla \mathbf{u}}{|\nabla \mathbf{u}|} - \gamma \mathbf{r} \tag{3}$$

coincides with the steady state of the gradient flow of the original Chan-Vese active contours [5], [3]. So minimizing eq. (1) is equivalent to finding an optimal contour which optimizes the original Chan-Vese energy function. Although the energy in eq. 1 does not have a unique global minimiser, a well defined minimiser can be found within the interval $[0, 1]^n$:

$$\mathbf{u}^* = \underset{\mathbf{u} \in [0,1]^n}{\arg\min} |\nabla \mathbf{u}| + \gamma \langle \mathbf{u}, \mathbf{r} \rangle \tag{4}$$

Note that this results in a minimiser which values are between 0 and 1. It is however desirable to have a segmentation result where the values of a minimiser are constrained to $(0, 1)$, i.e. a pixel belongs to a segment or not. Therefore $\mathbf{u}^*$ is tresholded, i.e.

$$\Phi_\alpha(u^*(x)) = \begin{cases} 1 & \text{if } u^*(x) > \alpha \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

with a predefined $\alpha \in [0, 1]$. In [9] it is shown that $\Phi_\alpha(\mathbf{u}^*)$ is a global minimiser for the energy in eq. (1) and by extension for the energy function of the original Chan-Vese active contour model. In [8] the convex energy function in eq. (1) was generalized in order to incorporate edge information:

$$E[\mathbf{u}] = |\nabla \mathbf{u}|_{\mathbf{g}} + \gamma \langle \mathbf{u}, \mathbf{r} \rangle \tag{6}$$

where $\mathbf{g}$ is the result of an edge detector, e.g. $\mathbf{g} = \frac{1}{1+|\nabla \mathbf{f}|}$. The active contour minimizing this energy function can be seen as a combination of edge based snake active contours [1] and the region based Chan-Vese active contours [3].

III. Multimodal foreground and background

Chan-Vese active contours assume that both the intensity of foreground and background have a unimodal probability distribution, preferably with a small variance. This assumption is reasonable if working on a local neighbourhood, i.e. a band around a level-set, but is generally not true when the complete image is taken into account, which is done with the convex energy active contours. Several approaches have been proposed to work around this constraint. Mao et al. adjusted the convex formulation to allow gradual changes of the objects intensity [10]. This however does not work if the foreground or background consists of abruptly changing intensities, e.g. if some part in the image has a specific texture. A solution for the segmentation of objects with texture was proposed in [11]. This of course does not solve the problem of having a background consisting of multiple objects, each with different expected intensity. In [12] Shang et al. proposed to combine level-sets with foreground and background intensity modelling using the Gaussian mixture model (GMM). We will generalize this approach and embed it in a convex formulation.

Instead of partitioning the image in two classes, i.e. foreground and background, we assume that the image is composed of multiple classes. For example k classes correspond to the foreground and l classes correspond to the background. For each of these classes we model the intensity. Using a Bayesian classifier, we can classify each pixel as belonging to a specific foreground and background class:

$$\begin{aligned} c_f^\dagger(x) &= \underset{i \in [1,k]}{\arg\max} P(c_f(x) = i) P(f(x) \mid c_f(x) = i) \\ c_b^\dagger(x) &= \underset{j \in [1,l]}{\arg\max} P(c_b(x) = j) P(f(x) \mid c_b(x) = j) \end{aligned} \tag{7}$$

Where $c_f(x) = i$ means that pixel x belongs to foreground class i. To simplify notations, we define

$$\begin{aligned} p_f(x, i) &= P(c_f(x) = i) P(f(x) \mid c_f(x) = i) \\ p_b(x, j) &= P(c_b(x) = j) P(f(x) \mid c_b(x) = j) \end{aligned} \tag{8}$$

Give the classifications in eq. (7), we can modify the Chan-Vese energy function in to the following term:

$$E[\mathbf{u}] = \sum_{t=1}^{mn} g(t) \nabla \mathbf{u}(t) - \sum_{t \in \Omega^+} p_f(t, c_f^\dagger(t)) - \sum_{t \in \Omega^-} p_b(t, c_b^\dagger(t)) \tag{9}$$

where $g(t)$ corresponds to the edge strength at pixel t and Ω^+ is the set of all pixels belonging to the found segment, Ω^- is its inverse, i.e. the set corresponding to all background pixels. This energy term tries to maximize the probability instead of minimizing the variance of the foreground and background. A

well defined global optimizer for the convex energy function in eq. (9) is defined by:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}\in[0,1]^n} |\nabla \mathbf{u}|_{\mathbf{g}} + \gamma\langle \mathbf{u}, \mathbf{p}\rangle \tag{10}$$

with

$$p(t) = -\max_{i\in[1,k]} p_f(t,i) + \max_{j\in[1,l]} p_b(t,j) \tag{11}$$

IV. Optimization

Due to the convexity of the energy function in eq. (9), a wide range of minimisers can be used to find an optimal contour $\mathbf{u}^\dagger$. One such optimization method is Split Bregman optimization, which is an efficient optimization technique for solving L1-regularized problems [13], [6]. In order to find a contour which minimizes eq. (9), the Split Bregman method will "de-couple" the L1 and the inner product, by introducing a new variable $\mathbf{d}$ and by putting constraints on the problem. This results in the following optimization problem:

$$(\mathbf{u}^\dagger, \mathbf{d}^\dagger) = |\mathbf{d}|_{\mathbf{g}} + \gamma\langle \mathbf{u}, \mathbf{p}\rangle \text{ such that } \mathbf{d} = \nabla \mathbf{u} \tag{12}$$

This optimization problem can be converted to an unconstrained problem by adding a quadratic penalty function, i.e.

$$(\mathbf{u}^\dagger, \mathbf{d}^\dagger) = |\mathbf{d}|_{\mathbf{g}} + \gamma\langle \mathbf{u}, \mathbf{p}\rangle + \frac{\lambda}{2}|\mathbf{d} - \nabla \mathbf{u}|_2^2 \tag{13}$$

Note that the quadratic penalty function only approximates the constraint $\mathbf{d} = \nabla \mathbf{u}$. By using a Bregman iteration technique [13], this constraint can be enforced exactly in an efficient way. In the Bregman iteration technique an extra vector, $\mathbf{b_k}$ is subtracted from the penalty function. Then the following two unconstrained steps are iteratively solved.

$$\begin{aligned}(\mathbf{u_{k+1}}, \mathbf{d_{k+1}}) &= \arg\min_{\mathbf{u_k},\mathbf{d_k}} |\mathbf{d_k}|_{\mathbf{g}} + \gamma\langle \mathbf{u_k}, \mathbf{p}\rangle \\ &\quad + \frac{\lambda}{2}|\mathbf{d_k} - \nabla \mathbf{u_k} - \mathbf{b_k}|_2^2 \qquad (14)\\ \mathbf{b_{k+1}} &= \mathbf{b_k} + \nabla \mathbf{u_{k+1}} - \mathbf{d_{k+1}} \qquad (15)\end{aligned}$$

The first step requires optimizing for two different vectors, $\mathbf{u}$ and $\mathbf{d}$. We approximate these optimal vectors by optimizing eq. (14) for $\mathbf{u}$ and $\mathbf{d}$ independently:

$$\begin{aligned}\mathbf{u_{k+1}} &= \arg\min_{\mathbf{u_k}} \gamma\langle \mathbf{u_k}, \mathbf{p}\rangle + \frac{\lambda}{2}|\mathbf{d_k} - \nabla \mathbf{u_k} - \mathbf{b_k}|_2^2 \\ \mathbf{d_{k+1}} &= \arg\min_{\mathbf{d_k}} |\mathbf{d_k}|_{\mathbf{g}} + \frac{\lambda}{2}|\mathbf{d_k} - \nabla \mathbf{u_{k+1}} - \mathbf{b_k}|_2^2\end{aligned} \tag{16}$$

The first problem can be optimized by solving a set of Euler-Lagrange equations. For each element $u(i)$ of the optimal $\mathbf{u}$ the following optimality condition should be satisfied:

$$\nabla^2 u(t) = \frac{\gamma}{\lambda} p(t) + \nabla \cdot (\mathbf{d}(t) - \mathbf{b}(t)) \tag{17}$$

Note that this system of equations can be written as $A\mathbf{u} = \mathbf{w}$. In [6] they proposed to solve this linear system using the iterative Gauss-Seidel method. In order to guarantee the convergence of this method, A should be strictly diagonally dominant or should be positive semi definite. Unfortunately A is neither. Instead we will optimize eq. (17) using the iterative conjugate residual method, which is a Krylov subspace method for which convergence is guaranteed if A is Hermitian [14].

The solution of eq. (17) is unconstrained, i.e. $u(i)$ does not have to lie in the interval $[0, 1]$. Note that minimizing eq. (16) for $u(i)$, i.e. all other elements of $\mathbf{u}$ remain constant, is equivalent to minimize a quadratic function. If $u(i) \notin [0, 1]$ then the constrained optimum is either 0 or 1, since a quadratic function is monotonic in an interval which does not contain its extremum. So the constrained optimum is given by:

$$u^*(i) = \max\Big(\min\big(u(i), 1\big), 0\Big) \tag{18}$$

In order to calculate an optimal $\mathbf{d_k}$ in eq. (16), a closed form solution can be calculated using the shrinking operator, i.e.

$$\mathbf{d_{k+1}}(t) = \text{shrink}\Big(\nabla \mathbf{u}(t) + \mathbf{b_k}, g(t), \lambda\Big) \tag{19}$$

where

$$\text{shrink}(\tau, \theta, \lambda) = \begin{cases} 0 & \text{if } |\tau| \le \frac{\theta}{\lambda} \\ \tau - \frac{\theta}{\lambda}\,\text{sgn}(\tau) & \text{otherwise}\end{cases} \tag{20}$$

In algorithm 1 we give an overview in pseudo code of the complete optimization algorithm. The CR function solves eq. (17) using the conjugate residual method, given the parameters $\mathbf{b_k}, \mathbf{d_k}$ and $\mathbf{p_k}$.

Algorithm 1: Split Bregman for active contour segmentation

```
while |u*_{k+1} − u*_k|_2 > ε do
    u_{k+1} = CR(b_k, d_k, p_k)
    u*_{k+1} = max(min(u_{k+1}, 1), 0)
    d_{k+1} = shrink(∇u*_{k+1} + b_k, g, λ)
    b_{k+1} = b_k + ∇u*_{k+1} − d_{k+1}
end
s_k = φ_α(u*_{k+1})
```

V. Results

Both the convex Chan-Vese as the proposed method requires some prior knowledge. The Chan-Vese active contours need an expected mean of foreground and background, i.e. μ_f and μ_b in eq. (1). To estimate these values from the images, we first threshold the image using Otsu thresholding [15]. The expected means are set based on foreground and background segments resulting from this thresholding.

The proposed method assumes that both foreground as background can consist of multiple classes. For all these classes the method requires prior knowledge about the probability density function of the intensity of a pixel belonging to the specific class. For a wide range of applications the possible foreground and background objects, i.e. classes, are known in advance, e.g. in an MRI scan of the brains you can expect white matter, gray matter as foreground objects and bone, cerebral fluid and

muscle tissue as possible background classes. For such an application the intensity probability density functions for all classes could be calculated from a training set using a kernel density estimator.

For the remaining of this section we will assume that such a training set is not available. Therefore the probability density functions have to be estimated based on the image. To achieve this, the image histogram will be approximated using a Gaussian mixture model, i.e.

$$h(i) \approx \sum_{j=1}^{l} \alpha_j p_g(i; \mu_j, \sigma_j^2) \tag{21}$$

where α_i is a weighting parameter such that $\alpha_i \in [0, 1]$ and $\sum_{j=1}^{l} \alpha_j = 1$ and where $p_g(.; \mu, \sigma^2)$ represents the probability density function of a Gaussian distribution with mean equal to μ and σ the standard deviation. The number of Gaussian distributions used, l in eq. (21), could be defined by a human operator or could be iteratively incremented until the approximation of the Gaussian mixture model resembles the image histogram sufficiently. The probability parameters μ_i and σ_i can be calculated using the expectation maximization algorithm [12]. For the proposed segmentation technique we will assume that each Gaussian distribution in eq. (21) corresponds to the probability density function of a foreground or background class. Which classes are considered foreground depends on the application and should be defined by a human operator.

We will show the benefit of the proposed method based on two examples. As a first example we would like to segment bacteria in a microscopic image. On the top of Fig. 1 the segmentation result of the original convex Chan-Vese active contour is shown. As can be seen does the bimodal convex active contour results in erroneous segmentation since the light gray interior of the bacteria is discarded by the method. The result of our proposed method can be seen in the bottom part of Fig. 1. To get this result, the image histogram was approximated by a mixture of 3 Gaussian distributions. The Gaussian distribution with highest mean was considered to correspond to the probability distribution of the background. The two other Gaussian distributions are considered to represent two modalities of foreground pixels. As can be seen in Fig. 1 results the proposed method in correct segmentation of the complete bacteria and this without a training step for the required probability distributions.

As a second example we segment an MRI image. In Fig. 2 the segmentation result of the bimodal convex active contours is shown. To get a clear view of the segments we show the result in two different ways: one image where the outlines of the segments are imposed on the original image, and a second image, where the segments are shown in white on a black background. The first image shows a clear view on where the contour of the segment lies, whereas the second image helps making clear which segments are foreground and which parts background. The bimodal active contour finds a combination of white and gray matter as a foreground segment. The gray

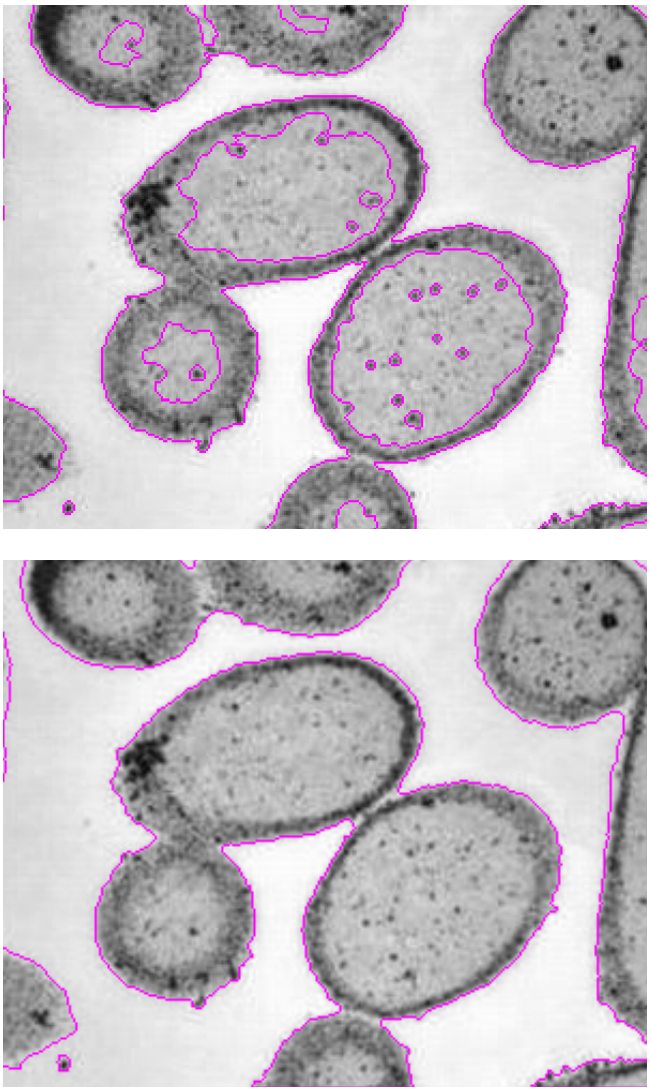

Figure 1. An example of segmentation of bacteria in a microscopic image. On the top, segmentation using the convex bimodal active contour is shown. In the bottom figure, segmentation using the proposed method is shown.

matter however is segmented erroneously, i.e. some parts of gray matter are considered foreground, whereas other parts are considered background. This might be solved by manually tuning μ_f in eq. (1), but even with tuned parameters, it is only possible to discriminate between dark background and bright foreground or vice versa. It is not possible to have a gray foreground and a background consisting of bright and dark parts. So using the bimodal active contours it is not possible to extract only the gray matter in one segmentation round.

The proposed method does not hamper from this drawback. In Fig. 3 the segmentation results using the proposed method are shown. For this example the image histogram has been modelled using four Gaussians: two Gaussians corresponding to white matter, one Gaussian for gray matter and one Gaussian for background and cerebral fluid. Fig. 3 shows three different segmentation results. The result depends on which classes are considered foreground: the top row corresponds

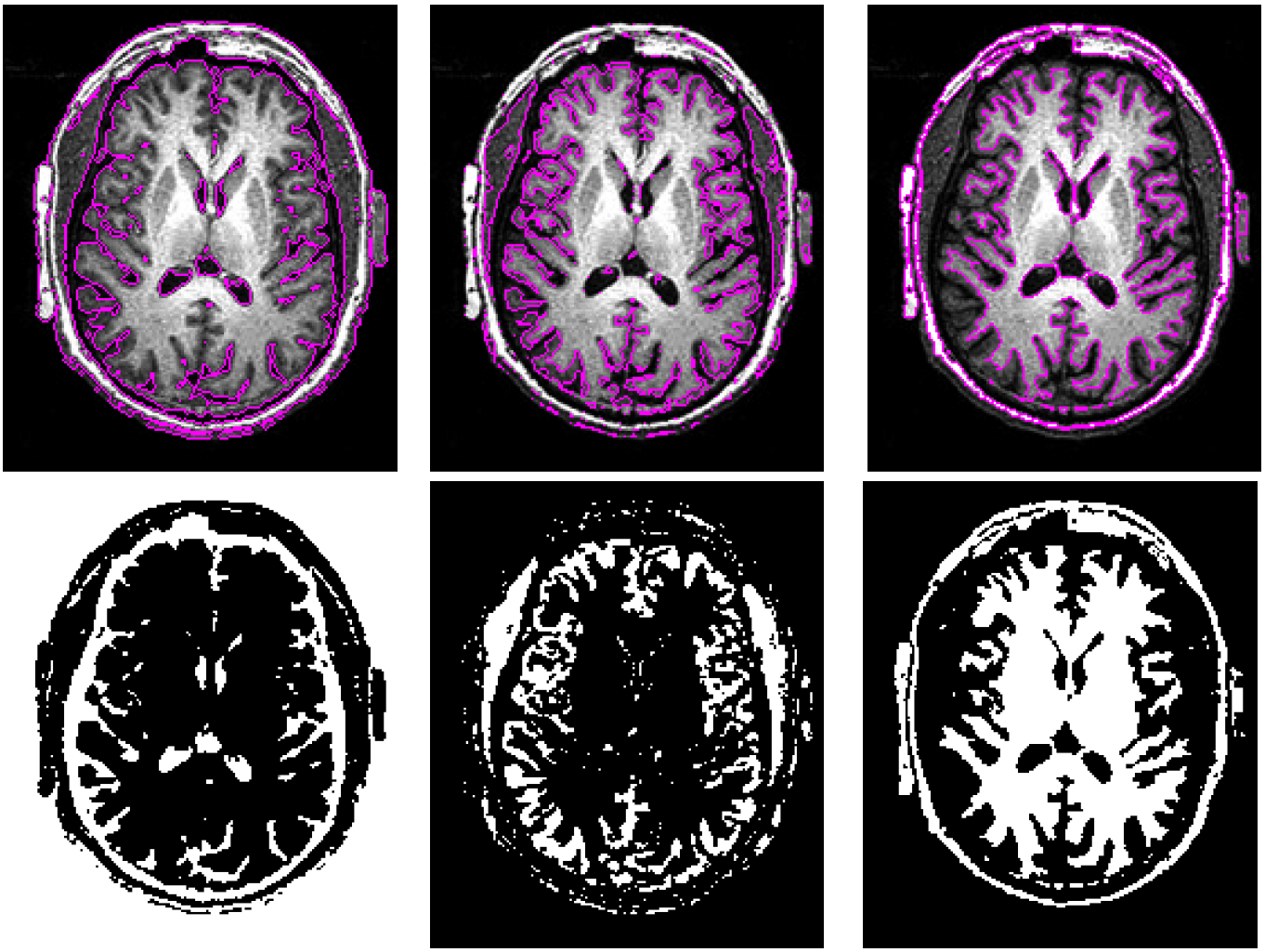

Figure 3. An example of segmentation in an MRI slice. Segmentation of the cerebral fluids can be seen in the left column, the middle column shows gray matter segments and the right column corresponds with segmentation of white matter.

	brain: white matter	bacteria
Chan-Vese	0.919	0.798
proposed method	0.952	0.9747

Table I
THE DICE COEFFICIENT BETWEEN THE SEGMENTATION RESULT AND MANUALLY DELINEATED GROUND TRUTH

with cerebral fluid, the middle row depicts gray matter and the bottom row corresponds to white matter.

In Table I a quantative comparison of the proposed method with the convex Chan-Vese segmentation is shown. The segmentation quality is compared with manual segmented ground truth using the Dice coefficient. The Dice coefficient is equal to one if the ground truth and the segmentation are identical and is zero if the ground truth segment has no pixels in common with the segmentation result.

The proposed optimization method was implemented in C and run on a computer with an Intel i7 Q720 1.6 GHz CPU with 4GB RAM. A single iteration takes about 60 ms for an image of dimension 512×512. The tested images both converged in less then 10 iterations. Note that the convex Chan-Vese active contours can be optimized using the same optimization method [6], so there is no computational difference between both methods.

VI. CONCLUSION

This paper proposes a new convex formulation for active contours. In contrast to convex active contour formulations found in literature, does the proposed formulation not assume a bright foreground and dark background or vice versa. Instead it assumes that both foreground and background consists of multiple classes, i.e. multiple modes in intensity. Using prior knowledge of these classes the constraint of bimodality has been removed. The required prior knowledge could be calculated from a training set or could be estimated based on the image itself. A suitable estimation method based on Gaussian mixture modelling has been proposed. Two real

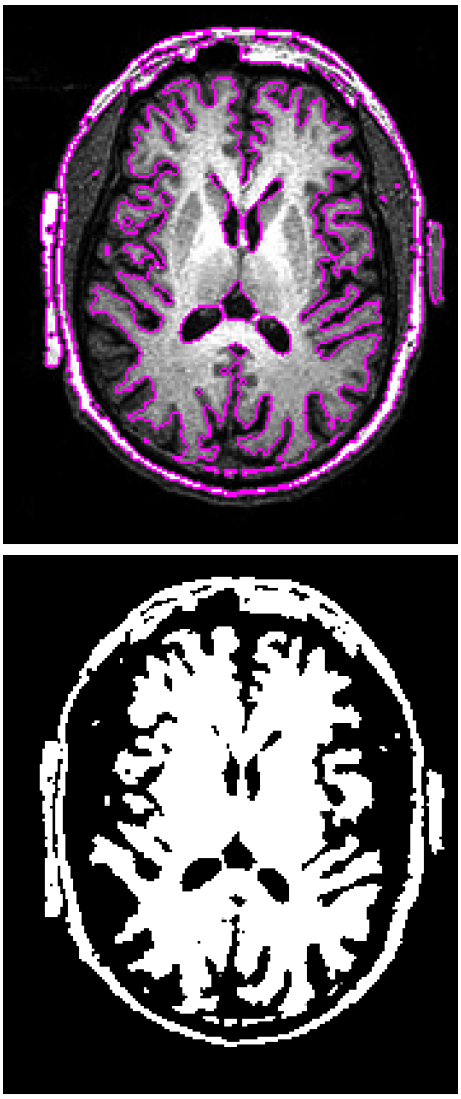

Figure 2. An example of segmentation in an MRI slice using the convex bimodal active contours.

applications showed the advantage of the proposed method over the original convex bimodal active contour segmentation.

REFERENCES

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International journal of computer vision*, pp. 321–331, 1988.

[2] M. Isard and A. Blake, *Active contours*. Springer, 1998.

[3] T. Chan and L. Vese, "An active contour model without edges," *Scale-Space Theories in Computer Vision*, vol. 1682, pp. 141–151, 1999.

[4] M.-A. Charmi, S. Derrode, and S. Ghorbel, "Fourier-based geometric shape prior for snakes," *Pattern Recognition Letters*, vol. 29, pp. 897–904, 2008.

[5] T. F. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *Siam Journal on Applied Mathematics*, vol. 66, no. 5, pp. 1632–1648, 2006.

[6] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split bregman method: Segmentation and surface reconstruction," *Journal of Scientific Computing*, vol. 45, no. 1-3, pp. 272–293, 2010.

[7] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, "Fast geodesic active contours," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1467–1475, 2001.

[8] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Thiran, and S. Osher, "Fast global minimization of the active contour/snake model," *Journal of Mathematical Imaging and Vision*, vol. 28, no. 2, pp. 151–167, 2007.

[9] X. Bresson and T. F. Chan, "Active contours based on chambolle's mean curvature motion," *2007 IEEE International Conference on Image Processing, Vols 1-7*, pp. 33–363 371, 2007.

[10] H. Mao, H. Liu, and P. Shi, "A convex neighbor-constrained active contour model for image segmentation," in *IEEE International Conference on Image Processing*, 2010.

[11] N. Houhou, J. Thiran, and X. Bresson, "Fast Texture Segmentation Based on Semi-local Region Descriptor and Active Contour," *Numerical Mathematics: Theory , Methods and Applications.*, vol. 2, no. 4, pp. 445–468, 2009.

[12] Y. Shang, R. Deklerck, E. Nyssen, A. Markova, J. de Mey, X. Yang, and K. Sun, "Vascular active contour for vessel tree segmentation," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 4, pp. 1023 –1032, april 2011.

[13] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *Siam Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[14] Y. Saad, *Iterative methods for sparse linear systems*, 2nd ed. Philadelphia: SIAM, Year.

[15] N. Otsu, "Threshold selection method from gray-level histograms," *IEEE Transactions on Systems Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.